Assignment – 6.1

Question No.1

|  |
| --- |
| boxplot(titanic3\_1\_$fare ~ titanic3\_1\_$pclass, ylim = c(0,300), xlab = "class", ylab = "fare", Col = c("blue")) |
|  |
| |  | | --- | |  | |

The box plot shows that there is difference in the fare by class. The 1st class has a higher median fare compared to the 2nd and 3rd classes, and it also shows a great dispersion around the median. The inter-quartile range of 1st class also shows that there is a wide spread variation in the fare charged to the passengers in that class, while the 3rd class shows the lowest variation in the fares charged to passengers. There are outliers in all the classes, who are charged quite higher than 3rd quartile + 1.5 times the inter-quartile range. The outliers are towards the higher end displaying that the passengers are charged exhorbitantly high fare compared to others.

Question No.2

|  |
| --- |
|  |
|  |
|  |

titanic\_gen\_class <- as.data.frame(cbind(titanic3\_1\_$pclass,titanic3\_1\_$sex))

> class(titanic\_gen\_class)

[1] "data.frame"

colnames(titanic\_gen\_class)[1] <- "pclass"

> colnames(titanic\_gen\_class)[2] <- "gender"

> head(titanic\_gen\_class)

pclass gender

1 1 female

2 1 male

3 1 female

4 1 male

5 1 female

6 1 male

> table(titanic\_gen\_class)

gender

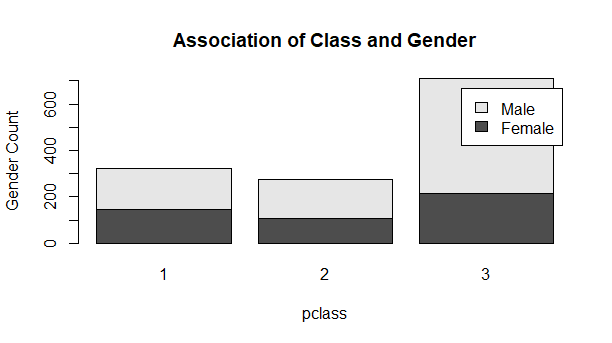
pclass female male

1 144 179

2 106 171

3 216 493

> barplot(t(table(titanic\_gen\_class)),xlab = "pclass", ylab = "Gender Count", main = "Association of Class and Gender",legend = c("Female", "Male"))



The stacked barplot visualizes that definitely the proportion of male passengers compared to female passengers is highest in those travelling by 3rd class. Unless we test statistically for the significance of the association between the gender and the class through a cross tab and the chi square test, we cannot confidently conclude that given that a person is a male there are higher chances that he would travel in a 3rd class. The first two classes do not show any great odds in favour of the males for one to be able to predict that given a gender there preference would be more for a particular class